Machine Learning in Materials Processing & Characterization

Philipp Pelz[1]

[1]Materials Science and Engineering,

Corresponding author: Philipp Pelz,

Abstract

This course provides students with essential skills and practical knowledge to harness machine learning techniques for accelerating materials discovery and design. Specifically tailored for students interested in the new BSc program "KI-Materialtechnologie"/AI for materials technology", it provides hands-on experience with core and advanced machine learning methods—including neural networks, optimization strategies, and generative modelling—to tackle real-world materials science problems. The course focuses on experimental data: microstructures, images, spectra, and processing parameters, connecting the messy, nonlinear world of processing and characterization signals with machine learning tools.

Plain Language Summary

This course teaches how to apply machine learning to materials science problems, focusing on experimental data from characterization techniques (microscopy, spectroscopy) and processing parameters. Students learn to build ML pipelines for microstructure classification, process prediction, and spectral analysis, with emphasis on understanding the physics of data formation and avoiding common pitfalls in experimental ML workflows.

1 Machine Learning in Materials Processing & Characterization 4th Semester – 5 ECTS, 2h lecture + 2h exercises per week

1.1 Synergy Map

- **This course**: What ML can do with experimental data: microstructures, images, spectra, processing parameters.
- **Parallel ML intro course**: Teaches generic ML algorithms and image processing foundations (skimage, Fourier, wavelets, SVMs, Bayes classifiers).
- **"Materials Genomics" course**: Focuses on materials databases, descriptors, crystal graph representations, DFT data, high-throughput workflows, surrogate models.

1.2 Week-by-Week Curriculum (14 weeks)

1.2.1 Unit I — Foundations: From Materials Signals to Machine Learning (Weeks 1–3)

1.2.1.1 Week 1 – What makes materials data special?

- Types of data: micrographs, EBSD, EDS, EELS, XRD, process logs, thermal profiles, deformation curves.
- PSPP (Processing–Structure–Property–Performance) as a data graph.
- Why vision-based ML and time-series ML are central to processing & characterization.

1.2.1.2 Week 2 – Image formation & the physics of data

- How characterization creates data: resolution, contrast mechanisms, artifacts.
- Fourier optics intuition for students with their ML-intro foundations.
- Sampling, aliasing, denoising as model-based priors.

1.2.1.3 Week 3 – Experimental data quality & ML-readiness

- Annotation, segmentation, inter-annotator variance.
- Train/test leakage in materials workflows.

1.2.2 Unit II — ML for Microstructure: Vision & Representation (Weeks 4–6)

1.2.2.1 Week 4 – Classical microstructure quantification & its ML extension

- Grain size, phase fractions, orientation maps, lineal intercepts.

- From hand-crafted features → learned representations.

1.2.2.2 Week 5 – Convolutional Neural Networks for microstructure classification

- CNN filters as microstructure interpreters.
- Example tasks: grain-boundary segmentation, precipitate detection, melt pool defects.

1.2.2.3 Week 6 – Transfer learning & data scarcity in materials characterization

- How to train a model with 200 images instead of 200k.
- Representations from ImageNet vs self-supervised pretraining on microstructures.

1.2.3 Unit III — ML in Processing: Time-Series, Optimization, Thermal/Mechanical Data (Weeks 7–9)

1.2.3.1 Week 7 – Process monitoring & time-series ML

- Process logs: temperature cycles, additive manufacturing melt pool monitoring, SPS, rolling, heat treatment.
- Hidden Markov models, ARIMA, random forest regressors, RNNs (light introduction).

1.2.3.2 Week 8 – Process → structure regression & uncertainty

- Gaussian Processes (synergy with Materials Genomics' surrogate models, but here linked to experimental data).
- Uncertainty as a tool for process design.

1.2.3.3 Week 9 – Inverse problems in processing

- ML-guided process maps (AM: laser power vs scan speed; metallurgy: TTT/CCT approximations).
- Physics-informed ML vs naive regression.

1.2.4 Unit IV — ML for Characterization Signals (Weeks 10–12)

1.2.4.1 Week 10 – Spectral data: ML for XRD, EELS, EDS

- Peak detection, denoising, background removal.
- Dimensionality reduction (PCA, NMF, ICA).

1.2.4.2 Week 11 – ML for microscopy automation

- Auto-focusing, drift correction, parameter selection.
- Vision-based defect detection in EBSD or TEM.

1.2.4.3 Week 12 – Multi-modal data fusion

- Combining images + spectra + process parameters.
- Early vs late fusion.

1.2.5 Unit V — Project + Reflection (Weeks 13–14)

1.2.5.1 Week 13 – Mini-project workshop

Projects could be:

- Predict microhardness from heat-treatment + microstructure images.
- Segment phases in SEM images.
- Detect porosity in AM melt pool images.
- Denoise EELS/XRD spectra.
- Build a process map using Gaussian Processes.

Students must show:

1. data prep → 2. model selection → 3. evaluation → 4. uncertainty → 5. interpretation.

1.2.5.2 Week 14 – Presentations + critical evaluation

- Focus on explainability (CAMs, SHAP for simple models).
- Reflect on why ML sometimes fails on materials data.
- Wrap-up: Where ML is genuinely changing materials characterization.

1.3 Learning Outcomes

Students completing this course should be able to:

- Interpret materials characterization and processing data in an ML-ready way.
- Build ML pipelines for microstructure classification, process prediction, and spectral analysis.
- Understand the physics of image/signal formation well enough to avoid "garbage in → garbage out".
- Evaluate uncertainty and biases in experimental ML models.
- Combine processing and characterization data for property prediction.
- Critically evaluate claims about ML in materials science. ##

1.4 Lab possibilities:

- **Lab:** Exploring real microscopy datasets; noise, metadata, units.
- **Lab:** Fourier & wavelet inspection of SEM/TEM/optical micrographs.
- **Lab:** Correct vs broken experimental ML pipelines; data-leak horror stories.
- **Lab:** Using scikit-image to extract features; PCA on microstructure descriptors.
- **Lab:** Fine-tuning a pretrained model on SEM/optical images.
- **Lab:** Predicting hardness from heat-treatment curves.
- **Lab:** GP on process parameters (e.g., cooling rate → microstructure metric).
- **Lab:** Building process maps using ML surrogate models.
- **Lab:** NMF decomposition of EELS datasets; automatic phase identification in XRD.
- **Lab:** Implementing a simple "AI autofocus" or EBSD pattern classifier.
- **Lab:** Fusing XRD + microstructure representations for property prediction.

References

Source: Article Notebook